

1 **CLAIMS**

2       1. A processor-readable medium comprising processor-executable  
3 instructions for personalizing karaoke, the processor-executable instructions  
4 comprising instructions for:

5           segmenting visual content to produce a plurality of sub-shots;  
6           segmenting music to produce a plurality of music sub-clips; and  
7           displaying at least some of the plurality of sub-shots as a background to  
8 lyrics associated with the plurality of music sub-clips.

9  
10       2. The processor-readable medium as recited in claim 1, additionally  
11 comprising instructions for:

12           shortening some of the plurality of sub-shots to a length of a corresponding  
13 music sub-clip from within the plurality of music sub-clips.

14  
15       3. The processor-readable medium as recited in claim 1, wherein  
16 segmenting the visual content comprises instructions for:

17           dividing a shot into two sub-shots at a maximum peak of a frame difference  
18 curve; and

19           repeating the dividing to result in sub-shots shorter than a maximum sub-  
20 shot length.

1           4.     The processor-readable medium as recited in claim 1, additionally  
2 comprising instructions for:

3                 filtering the plurality of sub-shots according to importance; and

4                 filtering the plurality of sub-shots according to quality.

5  
6           5.     The processor-readable medium as recited in claim 4, wherein  
7 filtering the plurality of sub-shots according to quality comprises instructions for:

8                 examining color entropy within each of the plurality of sub-shots for  
9 indications of diffusion of color; and

10                if color entropy is low, analyzing each of the plurality of sub-shots to detect  
11 motion more that a threshold indicating interest and less than a threshold indicating  
12 low camera and/or object movement;

13                selecting sub-shots having acceptable motion and/or color entropy scores.

14  
15           6.     The processor-readable medium as recited in claim 4, wherein  
16 filtering the plurality of sub-shots according to importance comprises instructions  
17 for:

18                evaluating frames within a sub-shot according to attention indices; and

19                averaging the attention indices for the frames to determine if the sub-shot  
20 should be included or excluded.

1           7.     The processor-readable medium as recited in claim 4, wherein  
2 filtering the sub-shots according to importance comprises instructions for:

3           analyzing for camera motion, for object motion and for specific objects  
4 within the sub-shots;

5           filtering the sub-shots according to the analysis.  
6

7           8.     The processor-readable medium as recited in claim 1, wherein the  
8 instructions for segmenting visual content segment video.  
9

10          9.     The processor-readable medium as recited in claim 8, additionally  
11 comprising instructions for:

12           selecting important sub-shots from within the plurality of sub-shots; and

13           selecting sub-shots such that they are uniformly distributed within the  
14 video.  
15

16          10.    The processor-readable medium as recited in claim 9, wherein  
17 selecting important sub-shots comprises instructions for:

18           evaluating color entropy, camera motion, object motion and object  
19 detection; and

20           selecting the important sub-shots based on the evaluation.  
21

22          11.    The processor-readable medium as recited in claim 9, wherein  
23 selecting uniformly distributed sub-shots comprises instructions for:

24           evaluating normalized entropy of the sub-shots along a time line of video  
25 from which the sub-shots were obtained.

1  
2       **12.**   The processor-readable medium as recited in claim 1, wherein the  
3 instructions for segmenting visual content includes instructions for assigning  
4 photographs to be sub-shots.  
5

6       **13.**   The processor-readable medium as recited in claim 12, wherein the  
7 instructions for assigning photographs includes instructions for:  
8       rejecting photographs having problems with quality; and  
9       rejecting photographs within a group of very similar photographs wherein a  
10 photo within the group has been selected.  
11

12       **14.**   The processor-readable medium as recited in claim 12, wherein the  
13 instructions for assigning photographs includes instructions for:  
14       converting at least one of the photographs to video.  
15

16       **15.**   The processor-readable medium as recited in claim 1, wherein the  
17 visual content comprises home video and photographs in digital formats.  
18

19       **16.**   The processor-readable medium as recited in claim 1, wherein  
20 segmenting the music comprises instructions for:  
21       establishing boundaries for the music sub-clips at beat positions within the  
22 music.  
23  
24  
25

1       **17.** The processor-readable medium as recited in claim 1, wherein  
2 segmenting music into the plurality of music sub-clips comprises instructions for  
3 bounding music sub-clip length according to:

4       minimum length =  $\min\{\max\{2 * \text{tempo}, 2\}, 4\}$  and

5       maximum length = minimum + 2.

6  
7       **18.** The processor-readable medium as recited in claim 1, wherein  
8 segmenting the music comprises instructions for:

9       establishing music sub-clips' length within a range of 3 to 5 seconds.

10  
11       **19.** The processor-readable medium as recited in claim 18, wherein  
12 segmenting the music comprises instructions for:

13       establishing boundaries for the music sub-clips at sentence breaks.

14  
15       **20.** The processor-readable medium as recited in claim 1, additionally  
16 comprising instructions for:

17       obtaining the lyrics from a file; and

18       coordinating delivery of the lyrics with the music using timing information  
19 contained within the file.

20  
21       **21.** A processor-readable medium as recited in claim 20, wherein  
22 obtaining the lyrics comprises instructions for sending the file over a network to a  
23 karaoke device as a part of a pay-for-play service.

1           **22.**   The processor-readable medium as recited in claim 1, additionally  
2 comprising instructions for:

3                 querying a database of songs by humming a portion of a desired song; and  
4                 selecting the desired song from among a number of possibilities suggested  
5 by an interface to the database.

6  
7           **23.**   A processor-readable medium comprising processor-executable  
8 instructions for providing lyrics for integration with music suitable for karaoke, the  
9 processor-executable instructions comprising instructions for:

10                receiving a request for a file associated with a specified song, wherein the  
11 file:

12                         associates each syllable contained within the lyrics with timing  
13 values; and

14                         associates each sentence contained within the lyrics with timing  
15 values; and

16                fulfilling the request for the file by sending the file associated with the  
17 specified song.

18  
19           **24.**   A processor-readable medium as recited in claim 23, wherein  
20 obtaining the lyrics comprises instructions for sending the file over a network to a  
21 karaoke device.

1           **25.**    A personalized karaoke device, comprising:  
2           a music analyzer configured to create music sub-clips of varying lengths  
3 according to a song;  
4           a visual content analyzer configured to define and select visual content sub-  
5 shots;  
6           a lyric formatter configured to time delivery of syllables of lyrics of the  
7 song; and  
8           a composer configured to assemble the music-sub clips with the visual  
9 content sub-shots, and configured to adjust length of the sub-shots to correspond to  
10 the music sub-clips, and configured to superimpose the syllables of the lyrics of  
11 the song over the sub-shots.

12  
13           **26.**    The personalized karaoke device of claim 25, wherein the music  
14 analyzer is configured to segment the song with a strong onset between each of the  
15 music sub-clips.

16  
17           **27.**    The personalized karaoke device of claim 25, wherein the music  
18 analyzer is configured to segment the song with a beat between each of the music  
19 sub-clips.

20  
21           **28.**    The personalized karaoke device of claim 25, wherein the music  
22 analyzer is configured to segment the song automatically into sub-clips, each  
23 having a duration that is a function of song tempo.

1           **29.**    The personalized karaoke device of claim 25, wherein the visual  
2 content analyzer is configured to segment video into sub-shots.

3  
4           **30.**    The personalized karaoke device of claim 25, wherein the visual  
5 content analyzer is configured to access folders of home video and photographs  
6 containing content from which the sub-shots are derived.

7  
8           **31.**    The personalized karaoke device of claim 25, wherein the visual  
9 content analyzer is configured to assemble still photographs, each of which is a  
10 sub-shot.

11  
12           **32.**    The personalized karaoke device of claim 25, wherein the visual  
13 content analyzer is configured to select from among sub-shots according to ranked  
14 importance, wherein importance is gauged by detection of color entropy, detection  
15 of object motion within the sub-shot, detection of camera motion during the sub-  
16 shot, and/or detection of a face within the sub-shot.

17  
18           **33.**    The personalized karaoke device of claim 25, wherein the visual  
19 content analyzer is configured to filter out sub-shots having low image quality as  
20 measured by low entropy and low motion intensity.

21  
22           **34.**    The personalized karaoke device of claim 25, wherein the visual  
23 content analyzer is configured to select sub-shots of greater importance consistent  
24 with creating a uniform distribution of the sub-shots over a runtime of a source  
25 video.



1  
2       **35.**    The personalized karaoke device of claim 25, wherein the visual  
3 content analyzer is configured to reject photographs of low quality by detecting  
4 over and under exposure, overly homogeneous images and blurred images.

5  
6       **36.**    The personalized karaoke device of claim 25, wherein the visual  
7 content analyzer is configured to organize photographs by date of exposure and by  
8 scene, thereby obtaining photographs having a relationship.

9  
10       **37.**   The personalized karaoke device of claim 37, wherein the visual  
11 content analyzer is configured to reject photographs which are members within a  
12 group of very similar photographs, wherein one of the group has already been  
13 selected.

14  
15       **38.**   The personalized karaoke device of claim 25, wherein the visual  
16 content analyzer is configured to:

17       detect an attention area within a photograph; and  
18       create a photo to video sub-shot based on the attention area, wherein the  
19 video includes panning and/or zooming.

20  
21       **39.**   The personalized karaoke device of claim 25, wherein the lyric  
22 formatter is configured to consume a file detailing timing of each syllable and each  
23 sentence of the lyrics.

1           **40.**    An apparatus, comprising:  
2           means for creating music sub-clips of varying lengths according to a song;  
3           means for defining and selecting visual content sub-shots;  
4           means for timing delivery of syllables of lyrics of the song; and  
5           means for assembling the music sub-clips with the visual content sub-shots,  
6           and to adjust length of the sub-shots to correspond to length of the music sub-clips,  
7           and to superimpose the syllables of the lyrics of the song over the sub-shots.  
8

9           **41.**    The apparatus of claim 40, wherein the means for defining and  
10          selecting visual content sub-shots is a video analyzer configured to segment video  
11          into sub-shots.  
12

13          **42.**    The apparatus of claim 40, wherein the means for defining and  
14          selecting visual content sub-shots is a video analyzer configured to access folders  
15          of home video and photographs containing content from which the sub-shots are  
16          derived.  
17

18          **43.**    The apparatus of claim 40, wherein the means for defining and  
19          selecting visual content sub-shots is a video analyzer configured for:  
20                  detecting an attention area within a photograph; and  
21                  creating a photo to video sub-shot based on the attention area, wherein the  
22          video includes panning and zooming.  
23  
24  
25

1           **44.**    The apparatus of claim 40, wherein the means for timing delivery of  
2 syllables of lyrics of the song is a lyric formatter configured for consuming a file  
3 detailing timing of each syllable and each sentence of the lyrics and for rendering  
4 the lyrics syllable by syllable.  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25